

ATGGCCCAAGCCCTGCCCTGGCTCCTGCTGTGGATGGGCGCGGGAG
TGCTGCCTGCCCACGGCACCCAGCACGGCATCCGGCTGCCCCCTGCG
CAGCGGCCTGGGGGGCGCCCCCTGGGGCTGCGGCTGCCCCGGGA
GACCGACGAAGAGCCCGAGGAGCCCGGCCGGAGGGGCGAGCTTTGT
GGAGATGGTGGACAACCTGAGGGGGCAAGTCGGGGCAGGGCTACTAC
GTGGAGATGACCGTGGGCAGCCCCCGCAGACGCTCAACATCCTGG
TGGATACAGGCAGCAGTAACTTTGCAGTGGGTGCTGCCCCCACCC
CTTCCTGCATCGCTACTACCAGAGGCAGCTGTCCAGCACATACCGGG
ACCTCCGGAAGGGTGTGTATGTGCCCTACACCCAGGGCAAGTGGGA
AGGGGAGCTGGGCACCGACCTGGTAAGCATCCCCCATGGCCCCAAC
GTCAGTGTGCGTGCCAACATTGCTGCCATCACTGAATCAGACAAGTT
CTTCATCAACGGGCTCCAACCTGGGAAGGCATCCTGGGGCTGGCCTATG
CTGAGATTGCCAGGCCTGACGACTCCCTGGAGCCTTTCTTTGACTCT
CTGGTAAAGCAGACCCACGTTCCCAACCTCTTCTCCCTGCAGCTTTG
TGGTGCTGGCTTCCCCCTCAACCAGTCTGAAGTGCTGGCCTCTGTGCG
GAGGGAGCATGATCATTGGAGGTATCGACCACTCGCTGTACACAGGC
AGTCTCTGGTATACACCCATCCGGCGGGAGTGGTATTATGAGGTGAT
CATTGTGCGGGTGGAGATCAATGGACAGGATCTGAAAATGGACTGCA
AGGAGTACAACCTATGACAAGAGCATTGTGGACAGTGGCACCACCAAC
CTTCGTTTGCCCAAGAAAGTGTTTGAAGCTGCAGTCAAATCCATCAAG
GCAGCCTCCTCCACGGAGAAGTTCCCTGATGGTTTCTGGCTAGGAGA
GCAGCTGGTGTGCTGGCAAGCAGGCACCACCCCTTGGAACATTTTCC
CAGTCATCTCACTCTACCTAATGGGTGAGGTTACCAACCAGTCCTTCC
GCATCACCATCCTTCCGCAGCAATACCTGCGGCCAGTGGAAGATGTG
GCCACGTCCCAAGACGACTGTTACAAGTTTGCCATCTCACAGTCATC
CACGGGCACTGTTATGGGAGCTGTTATCATGGAGGGCTTCTACGTTG
TCTTTGATCGGGCCCGAAAACGAATTGGCTTTGCTGTCAGCGCTTGC
CATGTGCACGATGAGTTCAGGACGGCAGCGGTGGAAGGCCCTTTTG
TCACCTTGGACATGGAAGACTGTGGCTACAACATTCCACAGACAGAT
GAGTCAACCCTCATGACCATAGCCTATGTCATGGCTGCCATCTGCGC
CCTCTTCATGCTGCCACTCTGCCTCATGGTGTGTCAGTGGCGCTGCC
TCCGCTGCCTGCGCCAGCAGCATGATGACTTTGCTGATGACATCTCC
CTGCTGAAG

FIG. 1A

CCATGCCGGCCCCCTCACAGCCCCGCCGGGAGCCCCGAGCCCCGCTGCCCAGG
 CTGGCCGCCGCSGTGCCGATGTAGCGGGCTCCGGATCCCAGCCTCTCCCCCT
 GCTCCCGTGCTCTGCGGATCTCCCCTGACCGCTCTCCACAGCCCCGGACCCG
 GGGGCTGGCCCCAGGGCCCTGCAGGCCCTGGCGTCCTGATGCCCCCAAGCT
 CCCTCTCCTGAGAAGCCACCAGCACCCAGACTTGGGGGCAGGCGCCA
 GGGACGGACGTGGGCCAGTGCGAGCCCAGAGGGCCCCGAAGGCCGGGGCC
 CACCATGGCCCAAGCCCTGCCCTGGCTCCTGCTGTGGATGGGCGCGGGAG
 TGCTGCCTGCCACGGCACCCAGCACGGCATCCGGCTGCCCTGCGCAGC
 GGCCTGGGGGGCGCCCCCCTGGGGCTGCGGCTGCCCCGGGAGACCGACG
 AAGAGCCCCGAGGAGCCCCGGCCGGAGGGGCAGCTTTGTGGAGATGGTGGAC
 AACCTGAGGGGGCAAGTCGGGGCAGGGCTACTACGTGGAGATGACCGTGGG
 CAGCCCCCGCAGACGCTCAACATCCTGGTGGATACAGGCAGCAGTAACTT
 TGCAGTGGGTGCTGCCCCCACCCTTCTGCTGCTACTACCAGAGGCA
 GCTGTCCAGCACATAACGGGACCTCCGGAAGGGTGTGTATGTGCCCTACAC
 CCAGGGCAAGTGGGAAGGGGAGCTGGGCACCGACCTGGTAAGCATCCCCC
 ATGGCCCCAACGTCACTGTGCGTGCCAACATTGCTGCCATCACTGAATCAGA
 CAAGTTCTTCATCAACGGCTCCAACCTGGGAAGGCATCCTGGGGCTGGCCTAT
 GCTGAGATTGCCAGGCCTGACGACTCCCTGGAGCCTTTCTTTGACTCTCTGG
 TAAAGCAGACCCACGTTCCCAACCTCTTCTCCCTGCAGCTTTGTGGTGTCTGG
 CTTCCCCCTCAACCAGTCTGAAGTGTGGCCTCTGTGCGGAGGGGAGCATGAT
 CATTGGAGGTATCGACCACTCGCTGTACACAGGCAGTCTCTGGTATACACCC
 ATCCGGCGGGAGTGGTATTATGAGGTGATCATTGTGCGGGTGGAGATCAAT
 GGACAGGATCTGAAAATGGACTGCAAGGAGTACAACTATGACAAGAGCATTG
 TGGACAGTGGCACCACCAACCTTCGTTTGCCCAAGAAAGTGTTTGAAGCTGC
 AGTCAAATCCATCAAGGCAGCCTCCTCCACGGAGAAGTTCCCTGATGGTTTC
 TGGCTAGGAGAGCAGCTGGTGTGCTGGCAAGCAGGCACCACCCCTTGGAAC
 ATTTTCCCAGTCATCTCACTCTACCTAATGGGTGAGGTTACCAACCAGTCCTT
 CCGCATCACCATCCTTCCGCAGCAATACCTGCGGCCAGTGGAAGATGTGGC
 CACGTCCCAAGACGACTGTTACAAGTTTGCCATCTCACAGTCATCCACGGGC
 ACTGTTATGGGAGCTGTTATCATGGAGGGCTTCTACGTTGTCTTTGATCGGG
 CCCGAAAACGAATTGGCTTTGCTGTGAGCGCTTGCCATGTGCACGATGAGTT
 CAGGACGGCAGCGGTGGAAGGCCCTTTTGTACCTTGGACATGGAAGACTG
 TGGCTACAACATTCCACAGACAGATGAGTCAACCCTCATGACCATAGCCTAT
 GTCATGGCTGCCATCTGCGCCCTCTTCATGCTGCCACTCTGCCTCATGGTGT
 GTCAGTGGCGCTGCCTCCGCTGCCTGCGCCAGCAGCATGATGACTTTGCTG
 ATGACATCTCCCTGCTGAAGTGAGGAGGCCCATGGGCAGAAGATAGAGATT
 CCCCTGGACCACACCTCCGTGGTTCACTTTGGTCACAAGTAGGAGACACAGA
 TGGCACCTGTGGCCAGAGCACCTCAGGACCCTCCCCACCCACCAAATGCCT
 CTGCCTTGATGGAGAAGGAAAAGGCTGGCAAGGTGGGTTCCAGGGACTGTA
 CCTGTAGGAAACAGAAAAGAGAAGAAAGAAGCACTCTGCTGGCGGGAATAC
 TCTTGGTCACCTCAAATTTAAGTCGGGAAATTCTGCTGCTTGAACTTCAGCC
 CTGAACCTTTGTCCACCATTCCTTTAAATTCTCCAACCCAAAGTATTCTTCTT
 TCTTAGTTTCAGAAGTACTGGCATCACACGCAGGTTACCTTGGCGTGTGTCC
 CTGTGGTACCCTGGCAGAGAAGAGACCAAGCTTGTTTCCCTGCTGGCCAAA
 GTCAGTAGGAGAGGATGCACAGTTTGCTATTTGCTTTAGAGACAGGGACTGT
 ATAAACAAGCCTAACATTGGTGCAAAGATTGCCTCTTGAATT

FIG. 1B

MAQALPWLLLWMGAGVLP AHGTQH GIRLPLRSG LGGAPLGLRL
PRETDEEPEEPGRRGSFVEMVDNL RGKSGQGYYVEMTVGSPP
QTLNILVDTGSSNFAVGAAPHPFLH RYYQRQLSSTYRDLRKGVY
VPYTQGKWE GELGTDLVSIPHGP NVTVRANIAAITESDKFFINGS
NWE GILGLAYAEIARPDDSLEPFFDSL VKQTHVPNLFSLQLCGAG
FPLNQSEVLASVGGSMIIGGIDHSLY TGSLWYTPIRREWYYEVIIV
RVEINGQDLKMDCKEYNYDKSIVDS GTTNLRLPKKVFEAAVKS IK
AASSTEKFPDGF WLGEQLVCWQAGTTPWNIFPVISLYLMGEVTN
QSFRITILPQQYLRPVEDVATSQDDCYKFAISQSSTGTVMGAVIM
EGFYVVFDRARKRIGFAVSACHVHDEFRTAAVEGPFVTLDMEDC
GYNIPQTDESTLMTIAYVMAAICALFMLPLCLMVCQWRCLRCLR
QQHDDFADDISLLK

FIG. 2A

ETDEEPEEPGRRGSFVEMVDNLRGKSGQGYVEMTVGSPPQT
LNILVDTGSSNFAVGAAPHPFLHRYYQRQLSSTYRDLRKGVYVP
YTQGKWEDELGTDLVSIPHGPNVTVRANIAAITESDKFFINGSNW
EGILGLAYAEIARPDDSLEPFFDSL VKQTHV PNLFS LQLCGAGFP
LNQSEVLASVGGSMIIGGIDHS LYTGSLWYTPIRREWYYEVIIVRV
EINGQDLKMDCKEYNYDKSIVDSGTTNLRLPKKVFEAAVKSIAA
SSTEKFPDGFWLGEQLVCWQAGTTPWNIFPVISLYLMGEVTNQ
SFRITILPQQYLRPVEDVATSQDDCYKFAISQSSTGTVMGAVIME
GFYVVFDRARKRIGFAVSACHVHDEFRTAAVEGPFVTLDMEDC
GYNIPQTDESTLMTIAYVMAAICALFMLPLCLMVCQWRCLRCLR
QQHDDFADDISLLK

FIG. 2B

MAQALPWLLLWMGAGVLP AHGTQH GIRLPLRSG LGGAPLGLRL
PRETDEEPEEPGRRGSFVEMVDNLRGKSGQGYVEMTVGSPP
QTLNILVDTGSSNFAVGAAPHPFLHRYYQRQLSSTYRDLRKG VY
VPYTQGKWE GELGTDLV SIPHGPNVTVRANIAAITESDKFFINGS
NWE GILGLAYAEIARPDDSLEPFFDSL VKQTHVPNLFSLQLCGAG
FPLNQSEVLASVGGSMIIGGIDHSLYTGSLWYTPIRREWYYEVIIV
RVEINGQDLKMDCKEYNYDKSIVDSGTTNLR LPPKKVFEEAAVKS I K
AASSTEKFPDGFWLGEQLVCWQAGTTPWNIFPVISLYLMGEVTN
QSFRITILPQQYLRPVEDVATSQDDCYKFAISQSSTGTVMGAVIM
EGFYVVFDRARKRIGFAVSACHVHDEFRTAAVEGPFVTLDMEDC
GYNIPQTDEEDYKDDDDK

FIG. 3A

ETDEEPEEPGRRGSFVEMVDNLRGKSGQGYVEMTVGSPPQT
LNILVDTGSSNFAVGAAPHPFLHRYYQRQLSSTYRDLRKG VYVP
YTQGKWE GELGTDLV SIPHGPNVTVRANIAAITESDKFFINGSNW
EGILGLAYAEIARPDDSLEPFFDSL VKQTHVPNLFSLQLCGAGFP
LNQSEVLASVGGSMIIGGIDHSLYTGSLWYTPIRREWYYEVIIVRV
EINGQDLKMDCKEYNYDKSIVDSGTTNLR LPPKKVFEEAAVKS I KAA
SSTEKFPDGFWLGEQLVCWQAGTTPWNIFPVISLYLMGEVTNQ
SFRITILPQQYLRPVEDVATSQDDCYKFAISQSSTGTVMGAVIME
GFYVVFDRARKRIGFAVSACHVHDEFRTAAVEGPFVTLDMEDC
GYNIPQTDEEDYKDDDDK

FIG. 3B

6/48

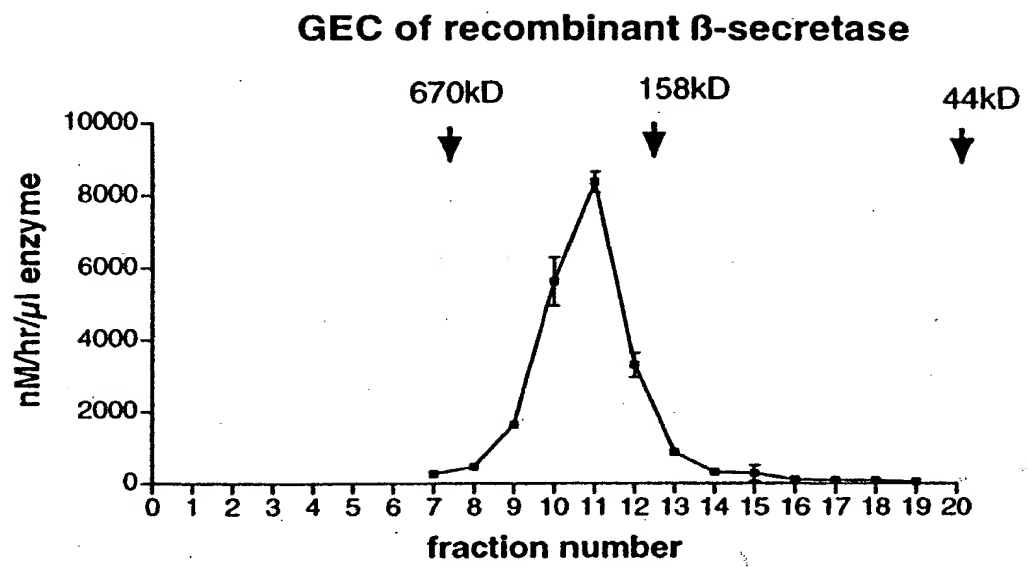


FIG. 4

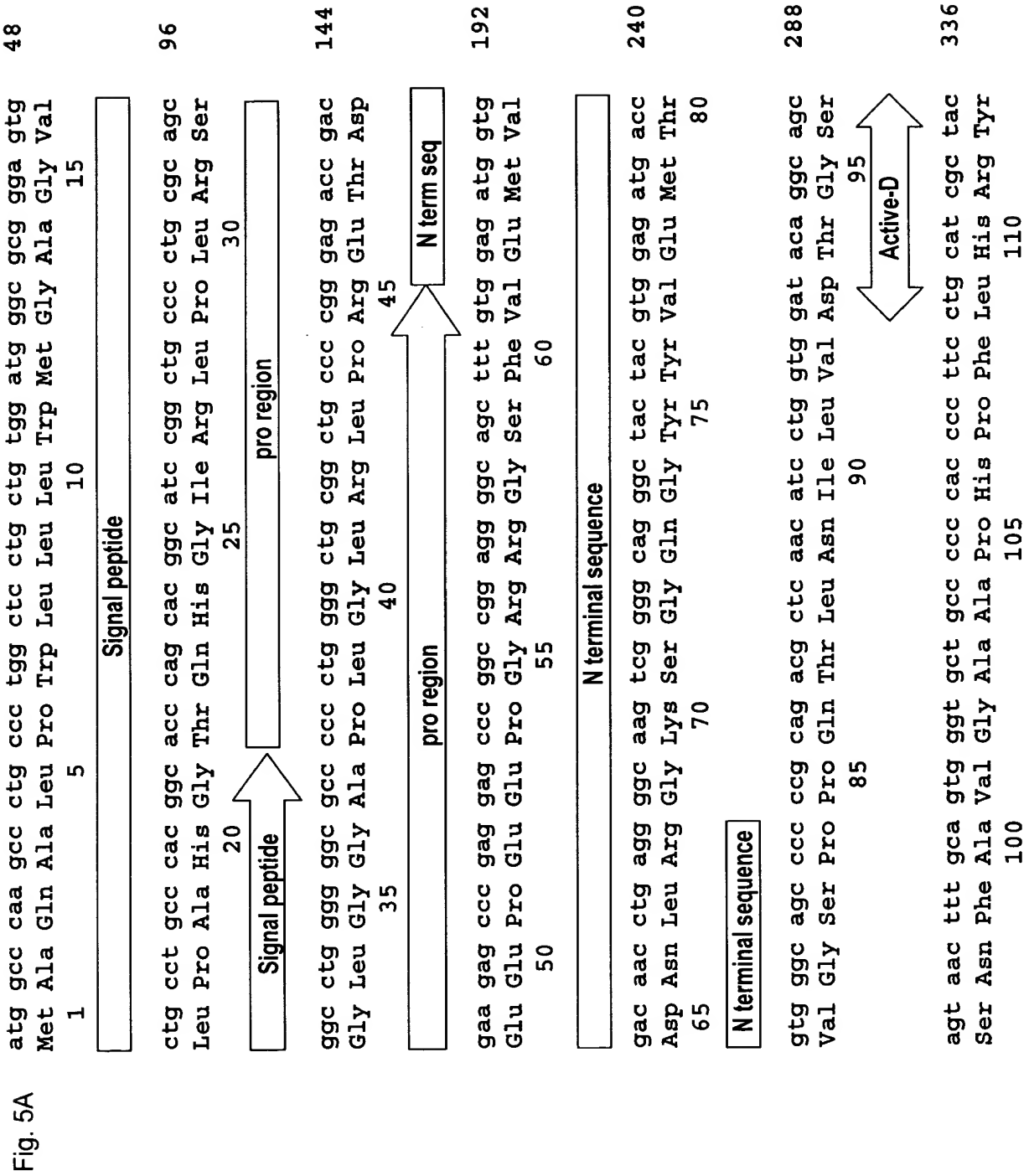


Fig. 5B

tac cag agg cag ctg tcc agc aca tac cgg gac ctc cgg aag ggt gtg	384
Tyr Gln Arg Gln Leu Ser Ser Thr Tyr Arg Asp Leu Arg Lys Gly Val	
115 120 125	
tat gtg ccc tac acc cag ggc aag tgg gaa ggg gag ctg ggc acc gac	432
Tyr Val Pro Tyr Thr Gln Gly Lys Trp Glu Gly Glu Leu Gly Thr Asp	
130 135 140	
ctg gta agc atc ccc cat ggc ccc aac gtc act gtg cgt gcc aac att	480
Leu Val Ser Ile Pro His Gly Pro Asn Val Thr Val Arg Ala Asn Ile	
145 150 155 160	
<div style="border: 1px solid black; padding: 2px; display: inline-block;">N-glycos</div>	
gct gcc atc act gaa tca gac aag ttc ttc atc aac ggc tcc aac tgg	528
Ala Ala Ile Thr Glu Ser Asp Lys Phe Phe Ile Asn Gly Ser Asn Trp	
165 170 175	
<div style="border: 1px solid black; padding: 2px; display: inline-block;">N-glycos</div>	
gaa ggc atc ctg ggg ctg gcc tat gct gag att gcc agg cct gac gac	576
Glu Gly Ile Leu Gly Leu Ala Tyr Ala Glu Ile Ala Arg Pro Asp Asp	
180 185 190	
tcc ctg gag cct ttc ttt gac tct ctg gta aag cag acc cac gtt ccc	624
Ser Leu Glu Pro Phe Phe Asp Ser Leu Val Lys Gln Thr His Val Pro	
195 200 205	
aac ctc ttc tcc ctg cag ctt tgt ggt gct ggc ttc ccc ctc aac cag	672
Asn Leu Phe Ser Leu Gln Leu Cys Gly Ala Gly Phe Pro Leu Asn Gln	
210 215 220	
<div style="border: 1px solid black; padding: 2px; display: inline-block;">N-glycos</div>	

Fig. 5C

tct gaa gtg ctg gcc tct gtc gga ggg agc atg atc att gga ggt atc Ser Glu Val Leu Ala Ser Val Gly Ser Met Ile Ile Gly Ile 225 230 235 240	720
N-gly	
gac cac tcg ctg tac aca ggc agt ctc tgg tat aca ccc atc cgg cgg Asp His Ser Leu Tyr Thr Gly Ser Leu Trp Tyr Thr Pro Ile Arg Arg 245 250 255	768
gag tgg tat tat gag gtg atc att gtg cgg gtg gag atc aat gga cag Glu Trp Tyr Tyr Glu Val Ile Ile Val Arg Val Glu Ile Asn Gly Gln 260 265 270	816
gat ctg aaa atg gac tgc aag gag tac aac tat gac aag agc att gtg Asp Leu Lys Met Asp Cys Lys Glu Tyr Asn Tyr Asp Lys Ser Ile Val 275 280 285	864
gac agt ggc acc acc aac ctt cgt ttg ccc aag aaa gtg ttt gaa gct Asp Ser Gly Thr Thr Asn Leu Arg Leu Pro Lys Lys Val Phe Glu Ala 290 295 300	912
<div> <div>Active-D</div> <div>↕</div> </div>	
gca gtc aaa tcc atc aag gca gcc tcc tcc acg gag aag ttc cct gat Ala Val Lys Ser Ile Lys Ala Ala Ser Ser Thr Glu Lys Phe Pro Asp 305 310 315 320	960
ggt ttc tgg cta gga gag cag ctg gtg tgc tgg caa gca ggc acc acc Gly Phe Trp Leu Gly Glu Gln Leu Val Cys Trp Gln Ala Gly Thr Thr 325 330 335	1008

Fig. 5D	cct tgg aac att ttc cca gtc atc tca ctc tac cta atg ggt gag gtt	1056
	Pro Trp Asn Ile Phe Pro Val Ile Ser Leu Tyr Leu Met Gly Glu Val	
	340 345 350	
	acc aac cag tcc ttc cgc atc acc atc ctt ccg cag caa tac ctg cgg	1104
	Thr Asn Gln Ser Phe Arg Ile Thr Ile Leu Pro Gln Gln Tyr Leu Arg	
	355 360 365	
	<div>N-glycos</div>	
	cca gtg gaa gat gtg gcc acg tcc caa gac gac tgt tac aag ttt gcc	1152
	Pro Val Glu Asp Val Ala Thr Ser Ser Gln Asp Asp Cys Tyr Lys Phe Ala	
	370 375 380	
	atc tca cag tca tcc acg gcc act gtt atg gga gct gtt atc atg gag	1200
	Ile Ser Gln Ser Ser Thr Gly Thr Val Met Gly Ala Val Ile Met Glu	
	385 390 395 400	
	ggc ttc tac gtt gtc ttt gat cgg gcc cga aaa cga att ggc ttt gct	1248
	Gly Phe Tyr Val Val Phe Asp Arg Ala Arg Lys Arg Ile Gly Phe Ala	
	405 410 415	
	<div></div>	
	gtc agc gct tgc cat gtg cac gat gag ttc agg acg gca gcg gtg gaa	1296
	Val Ser Ala Cys His Val His Asp Glu Phe Arg Thr Ala Ala Val Glu	
	420 425 430	
	<div>Internal peptide sequence</div>	

Fig. 5E

ggc cct ttt gtc acc ttg gac atg gaa gac tgt ggc tac aac att cca	1344
Gly Pro Phe Val Thr Leu Asp Met Glu Asp Cys Gly Tyr Asn Ile Pro	
435 440 445	
cag aca gat gag tca acc ctg acc ata gcc tat gtc atg gct gcc	1392
Gln Thr Asp Glu Ser Thr Leu Met Thr Ile Ala Tyr Val Met Ala Ala	
450 455 460	
Transmembrane	
atc tgc gcc ctg ttc atg ctg cca ctg tgc ctg atg gtg tgt cag tgg	1440
Ile Cys Ala Leu Phe Met Leu Pro Leu Cys Leu Met Val Cys Gln Trp	
465 470 475 480	
Transmembrane	
cgc tgc ctg cgc tgc ctg cgc cag cag cat gat gac ttt gct gat gac	1488
Arg Cys Leu Arg Cys Leu Arg Cys Gln Gln His Asp Asp Phe Ala Asp Asp	
485 490 495	
atc tcc ctg ctg aag tga	1506
Ile Ser Leu Leu Lys	
500	

12/48

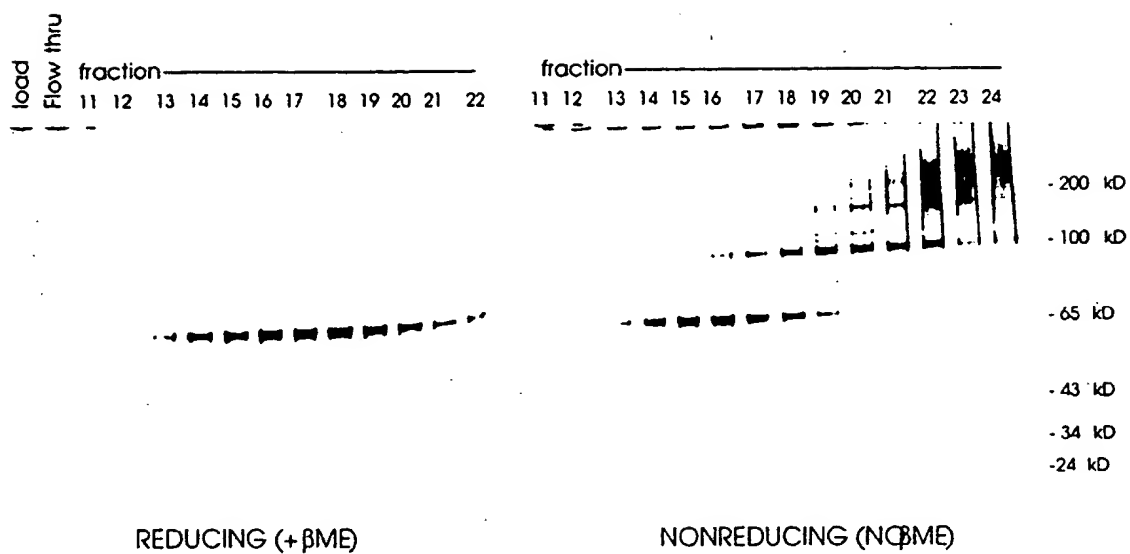


FIG. 6A

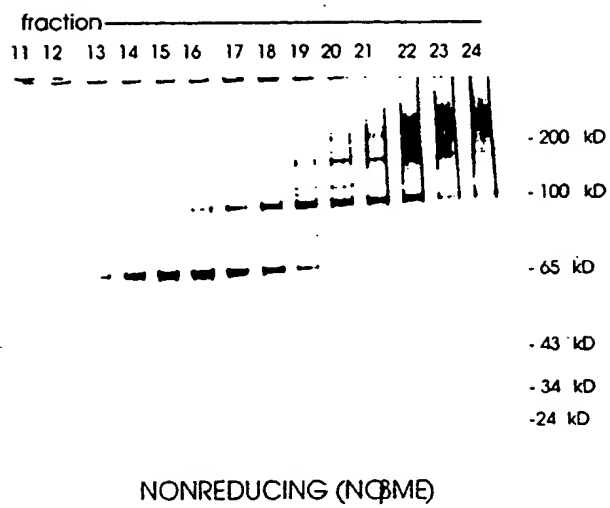


FIG. 6B

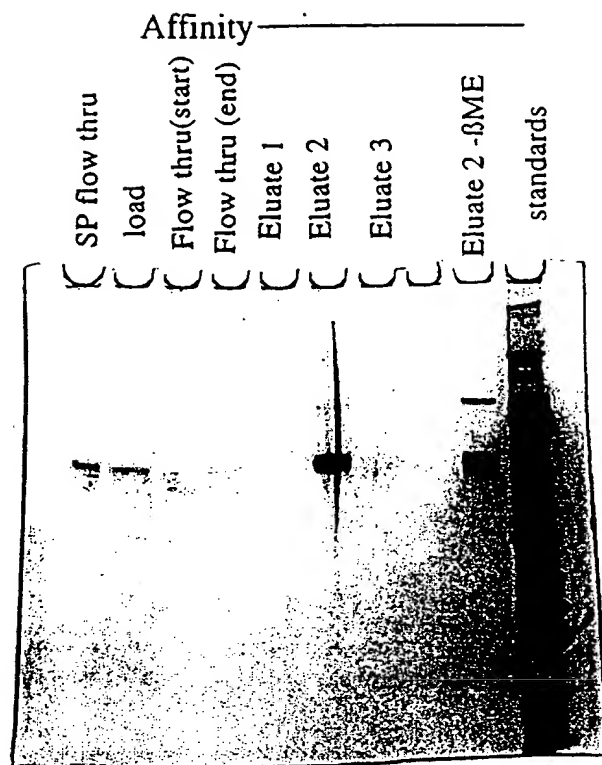


FIG. 7

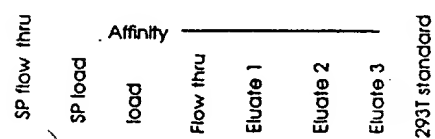


FIG. 8

14/48

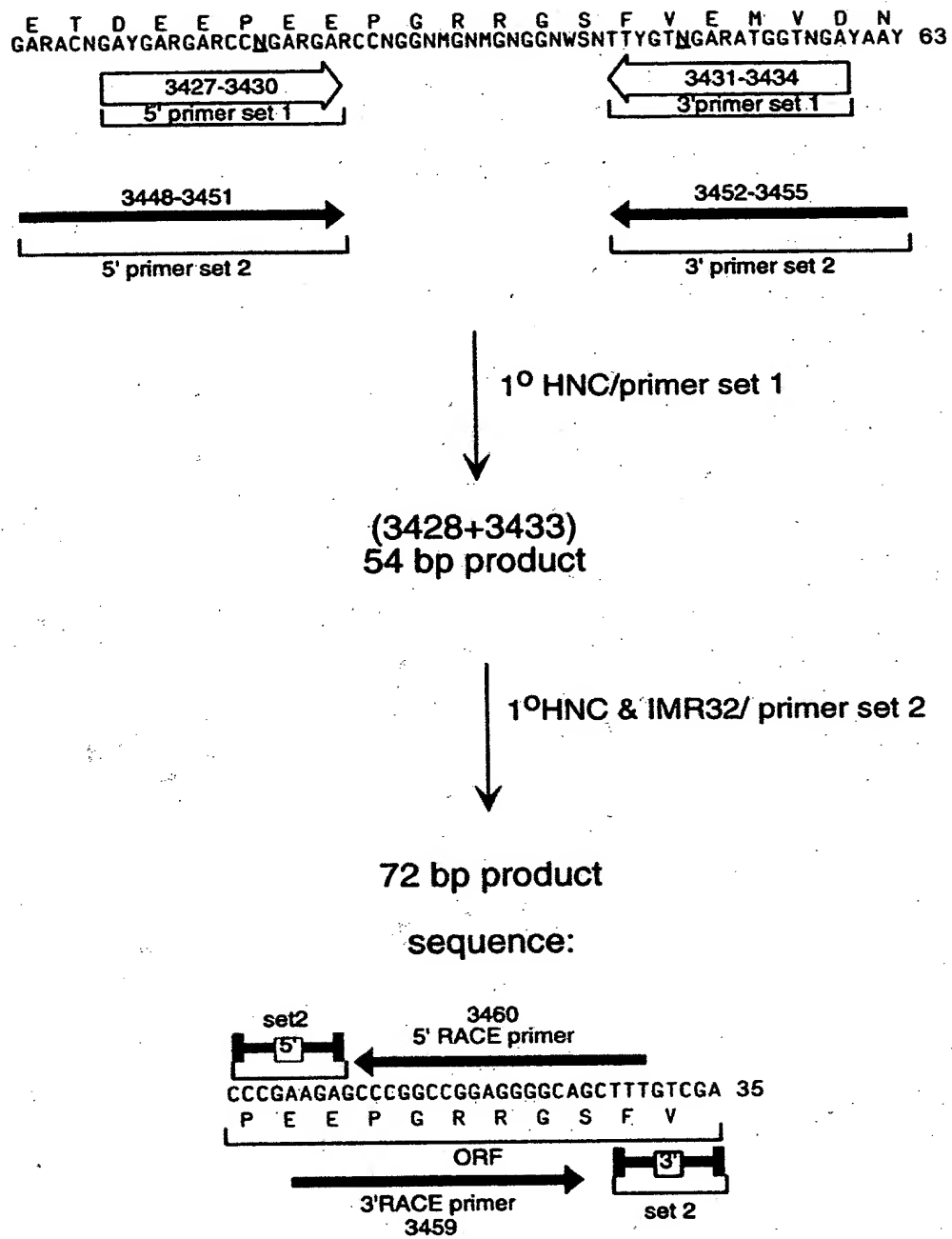


Fig. 9

	10	20	30	40	
Hump501prot	M A Q A L P W L L L W M G A G V L P A H G T Q H G I R L P L R S G L G G A P L G	40			
Musp501prot	M A P A L H W L L L W V G S G M L P A Q G T H L G I R L P L R S G L A G P P L G	40			
	50	60	70	80	
Hump501prot	L R L P R E T D E E P E E P G R R G S F V E M V D N L R G K S G Q G Y Y V E M T	80			
Musp501prot	L R L P R E T D E E S E E P G R R G S F V E M V D N L R G K S G Q G Y Y V E M T	80			
	90	100	110	120	
Hump501prot	V G S P P Q T L N I L V D T G S S N F A V G A A P H P F L H R Y Y Q R Q L S S T	120			
Musp501prot	V G S P P Q T L N I L V D T G S S N F A V G A A P H P F L H R Y Y Q R Q L S S T	120			
	130	140	150	160	
Hump501prot	Y R D L R K G V Y V P Y T Q G K W E G E L G T D L V S I P H G P N V T V R A N I	160			
Musp501prot	Y R D L R K G V Y V P Y T Q G K W E G E L G T D L V S I P H G P N V T V R A N I	160			
	170	180	190	200	
Hump501prot	A A I T E S D K F F I N G S N W E G I L G L A Y A E I A R P D D S L E P F F D S	200			
Musp501prot	A A I T E S D K F F I N G S N W E G I L G L A Y A E I A R P D D S L E P F F D S	200			
	210	220	230	240	
Hump501prot	L V K Q T H V P N I F S L Q L C G A G F P L N Q S E V L A S V G G S M I I G G I	240			
Musp501prot	L V K Q T H I P N I F S L Q L C G A G F P L N Q T E A L A S V G G S M I I G G I	240			
	250	260	270	280	
Hump501prot	D H S L Y T G S L W Y T P I R R E W Y Y E V I I V R V E I N G Q D L K M D C K E	280			
Musp501prot	D H S L Y T G S L W Y T P I R R E W Y Y E V I I V R V E I N G Q D L K M D C K E	280			
	290	300	310	320	
Hump501prot	Y N Y D K S I V D S G T T N L R L P K K V F E A A V K S I K A A S S T E K F P D	320			
Musp501prot	Y N Y D K S I V D S G T T N L R L P K K V F E A A V K S I K A A S S T E K F P D	320			
	330	340	350	360	
Hump501prot	G F W L G E Q L V C W Q A G T T P W N I F P V I S L Y L M G E V T N Q S F R I T	360			
Musp501prot	G F W L G E Q L V C W Q A G T T P W N I F P V I S L Y L M G E V T N Q S F R I T	360			
	370	380	390	400	
Hump501prot	I L P Q Q Y L R P V E D V A T S Q D D C Y K F A I S Q S S T G T V M G A V I M E	400			
Musp501prot	I L P Q Q Y L R P V E D V A T S Q D D C Y K F A V S Q S S T G T V M G A V I M E	400			
	410	420	430	440	
Hump501prot	G F Y V V F D R A R K R I G F A V S A C H V H D E F R T A A V E G P F V T L D M	440			
Musp501prot	G F Y V V F D R A R K R I G F A V S A C H V H D E F R T A A V E G P F V T A D M	440			
	450	460	470	480	
Hump501prot	E D C G Y N I P Q T D E S T L M T I A Y V M A A I C A L F M L P L C L M V C Q W	480			
Musp501prot	E D C G Y N I P Q T D E S T L M T I A Y V M A A I C A L F M L P L C L M V C Q W	480			
	490	500			
Hump501prot	R C L R C L R Q Q H D D F A D D I S L L K				501
Musp501prot	R C L R C L R H Q H D D F G D D I S L L K				501

FIG. 10

FIG. 10

CTGTTGGGCTCGCGGTTGAGGACAAACTCTTCGCGGTCTTTCCAGTACTCT
 TGGATCGGAAACCCGTCGGCCTCCGAACGGTACTCCGCCACCGAGGGACCT
 GAGCGAGTCCGCATCGACCGGATCGGAAAACCTCTCGACTGTTGGGGTGAG
 TACTCCCTCTCAAAAGCGGGCATGACTTCTGCGCTAAGATTGTCAGTTTCC
 AAAAACGAGGAGGATTTGATATTCACCTGGCCCCGCGGTGATGCCTTTGAGG
 GTGGCCGCGTCCATCTGGTCAGAAAAGACAATCTTTTTGTTGTCAAGCTTG
 AGGTGTGGCAGGCTTGAGATCTGGCCATACACTTGAGTGACAATGACATCC
 ACTTTGCCTTTCTCTCCACAGGTGTCCACTCCCAGGTCCAAGTGCAGGTCCG
 ACTCTAGACCC

FIG. 11A

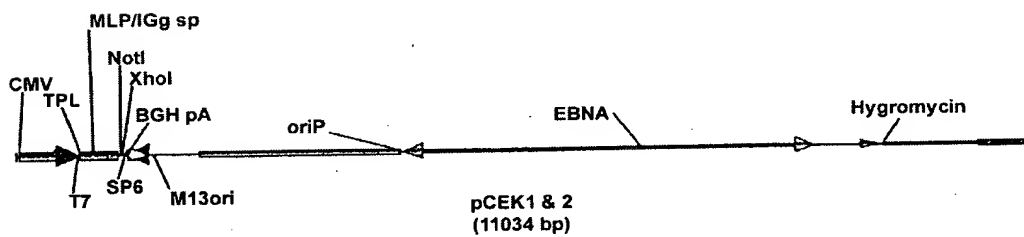


FIG. 11B

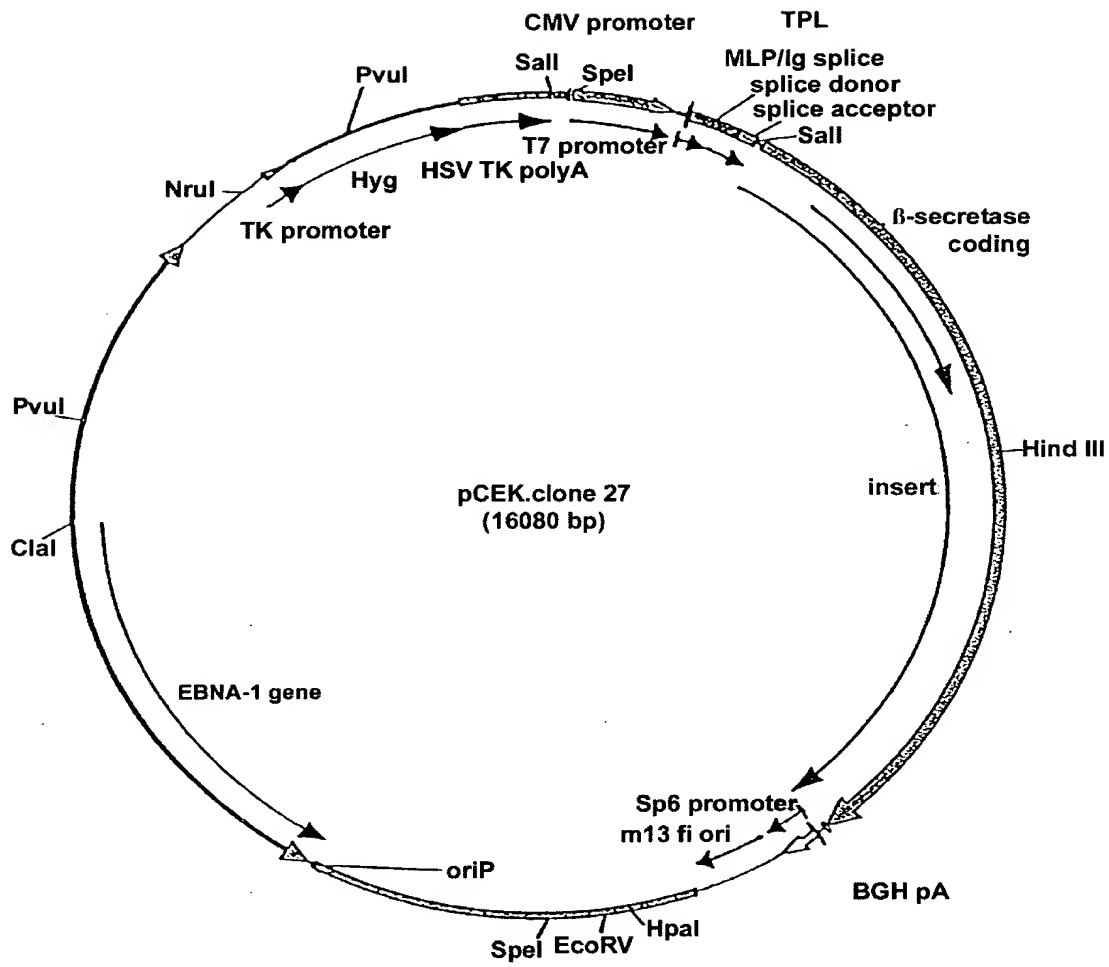


FIG. 12

ttctcatgtt tgacagctta tcatcgaga tccgggcaac gttgttgcac tgctgcaggc 60
gcagaactgg taggtatgga agatccgatg tacggggccag atatacgcggt tgacattgat 120
SpeI
tattgactag ttattaatag taatcaatta cgggggtcatt agttcatagc ccataatatgg 180
agttccgcgt tacataactt acggtaaatg gccgcctgg ctgaccgccc aacgaccccc 240
gcccatcgac gtcaataatg acgtatgttc ccatagtaac gccaataggg actttccatt 300
gacgtcaatg ggtggactat ttacggtaaa ctgcccactt ggcagtacat caagtgtatc 360
atatgccaag tacgccccct attgacgtca atgacggtaa atggcccgcc tggcattatg 420
cccagtaacat gaccttatgg gactttccta cttggcagta catctacgta ttagtcatcg 480
ctattaccat ggtgatgcgg ttttggcagt acatcaatgg gcgtggatag cggtttgact 540
cacgggggatt tccaagtctc caccattg acgtcaatgg gagtttgttt tggcaccaaa 600
atcaacggga ctttccaaa tgtcgtaaca actccgcccc attgacgcaa atgggcggtgta 660
ggcgtgtacg gtgggaggtc tatataagca gagctctctg gctaaactaga gaaccactg 720
cttactggct tatcgaaatt aatacgactc actataggga gacccaagct ctgttgggct 780

Figure 13B

cgcggttgag gacaaactct tcgcggtctt tccagtactc ttggatcga aacccgtcgg 840

 cctccgaacg gtactccgcc accgagggac ctgagcgagt cgcgcatcgac cggatcggaa 900
 splice donor
 aacctctga ctgttgggtt gagtactccc tctcaaaagc gggcatgact tctgcgctaa 960

 gattgtcagt ttccaaaaac gaggaggatt tgatatcac ctggccccgcg gtgatgcctt 1020

 tgagggtggc cgcgtccatc tggtcagaaa agacaatctt tttgttgtca agcttgaggt 1080

 gtggcaggct tgagatctgg ccatacactt gagtgaacaat gacatccact ttgcctttct 1140
 splice acceptor SalI
 ctccacaggt gtccactccc aggtccaact gcaggtcgac tctagaccgc gggaaattctg 1200

 cagatatcca tcacactggc cgcactcgtc ccagccccgc cgggagctg cgagccgcga 1260

 gctggattat ggtggcctga gcagccaacg cagccgcagg agcccggagc ccttgccccct 1320

 gcccgcgccg ccgcccgcg gggggaccag ggaagccgc accggccgc catgcccgc 1380

 cctccccagc ccgcccggag cccgcgcccg ctgcccaggc tggccgcccgc cgtgccgatg 1440

 tagcgggctc cggatcccag cctctccccct gctccccgtgc tctcgggatc tcccctgacc 1500

 gctctccaca gcccggacc gggggctggc ccaggggccct gcaggccctg gcgtcctgat 1560

 gcccccaagc tccctctcct gagaagccac cagcaccacc cagacttggg ggcaggcgcc 1620

Figure 13C

1677	aggacggac gtgggccagt gcgagcccag agggcccgaaggccggggcc cacc atg Met
	<u>1</u>
1725	gcc caa gcc ctg ccc tgg ctc ctg ctg tgg atg ggc ggc gga gtg ctg Ala Gln Ala Leu Pro Trp Leu Leu Trp Met Gly Ala Gly Val Leu
	5 10 15
1773	cct gcc cac ggc acc cag cac ggc atc cgg ctg ccc ctg cgc agc ggc Pro Ala His Gly Thr Gln His Gly Ile Arg Leu Pro Leu Arg Ser Gly
	20 25 30
1821	ctg ggg ggc gcc ccc ctg ggg ctg cgg ctg ccc cgg gag acc gac gaa Leu Gly Gly Ala Pro Leu Gly Leu Arg Leu Pro Arg Glu Thr Asp Glu
	35 40 45
1869	gag ccc gag gag ccc ggc cgg agg ggc agc ttt gtg gag atg gtg gac Glu Pro Glu Glu Pro Gly Arg Arg Gly Ser Phe Val Glu Met Val Asp
	50 55 60 65
1917	aac ctg agg ggc aag tcg ggg cag ggc tac tac gtg gag atg acc gtg Asn Leu Arg Gly Lys Ser Gly Gln Gly Tyr Tyr Val Glu Met Thr Val
	70 75 80
1965	ggc agc ccc ccg cag acg ctc aac atc ctg gtg gat aca ggc agc agt Gly Ser Pro Pro Gln Thr Leu Asn Ile Leu Val Asp Thr Gly Ser Ser
	85 90 95

Figure 13D

2013	aac ttt gca gtg ggt gct gcc ccc cac ccc ttc ctg cat cgc tac tac Asn Phe Ala Val Gly Ala Ala Pro His Pro Phe Leu His Arg Tyr Tyr	100 105 110
2061	cag agg cag ctg tcc agc aca tac cgg gac ctc cgg aag ggt gtg tat Gln Arg Gln Leu Ser Ser Thr Tyr Arg Asp Leu Arg Lys Gly Val Tyr	115 120 125
2109	gtg ccc tac acc cag ggc aag tgg gaa ggg gag ctg ggc acc gac ctg Val Pro Tyr Thr Gln Gly Lys Trp Glu Gly Glu Leu Gly Thr Asp Leu	130 135 140 145
2157	gta agc atc ccc cat ggc ccc aac gtc act gtg cgt gcc aac att gct Val Ser Ile Pro His Gly Pro Asn Val Thr Val Arg Ala Asn Ile Ala	150 155 160
2205	gcc atc act gaa tca gac aag ttc ttc atc aac ggc tcc aac tgg gaa Ala Ile Thr Glu Ser Asp Lys Phe Phe Ile Asn Gly Ser Asn Trp Glu	165 170 175
2253	ggc atc ctg ggg ctg gcc tat gct gag att gcc agg cct gac gac tcc Gly Ile Leu Gly Leu Ala Tyr Ala Glu Ile Ala Arg Pro Asp Asp Ser	180 185 190
2301	ctg gag cct ttc ttt gac tct ctg gta aag cag acc cac gtt ccc aac Leu Glu Pro Phe Phe Asp Ser Leu Val Lys Gln Thr His Val Pro Asn	195 200 205

Figure 13E

ctc ttc tcc ctg cag ctt tgt ggt gct ggc ttc ccc ctc aac cag tct Leu Phe Ser Leu Gln Leu Cys Gly Ala Gly Phe Pro Leu Asn Gln Ser 210 215 220 225	2349
gaa gtg ctg gcc tct gtc gga ggg agc atg atc att gga ggt atc gac Glu Val Leu Ala Ser Val Gly Ser Met Ile Ile Gly Gly Ile Asp 230 235 240	2397
cac tcg ctg tac aca ggc agt ctc tgg tat aca ccc atc cgg cgg gag His Ser Leu Tyr Thr Gly Ser Leu Trp Tyr Thr Pro Ile Arg Arg Glu 245 250 255	2445
tgg tat tat gag gtc atc att gtg cgg gtg gag atc aat gga cag gat Trp Tyr Tyr Glu Val Ile Ile Val Arg Val Glu Ile Asn Gly Gln Asp 260 265 270	2493
ctg aaa atg gac tgc aag gag tac aac tat gac aag agc att gtg gac Leu Lys Met Asp Cys Lys Glu Tyr Asn Tyr Asp Lys Ser Ile Val Asp 275 280 285	2541
agt ggc acc acc aac ctt cgt ttg ccc aag aaa gtg ttt gaa gct gca Ser Gly Thr Thr Asn Leu Arg Leu Pro Lys Lys Val Phe Glu Ala Ala 290 295 300 305	2589
gtc aaa tcc atc aag gca gcc tcc tcc acg gag aag ttc cct gat ggt Val Lys Ser Ile Lys Ala Ala Ser Ser Thr Glu Lys Phe Pro Asp Gly 310 315 320	2637

Figure 13F

2685	ttc tgg cta gga gag cag ctg gtg tgc tgg caa gca ggc acc acc cct	325	330	335
	Phe Trp Leu Gly Glu Gln Leu Val Cys Trp Gln Ala Gly Thr Thr Pro			
2733	tgg aac att ttc cca gtc atc tca ctc tac atg ggt gag gtt acc	340	345	350
	Trp Asn Ile Phe Pro Val Ile Ser Leu Tyr Leu Met Gly Glu Val Thr			
2781	aac cag tcc ttc cgc atc acc atc ctt ccg cag caa tac ctg cgg cca	355	360	365
	Asn Gln Ser Phe Arg Ile Thr Ile Leu Pro Gln Gln Tyr Leu Arg Pro			
2829	gtg gaa gat gtg gcc acg tcc caa gac gac tgt tac aag ttt gcc atc	370	375	380
	Val Glu Asp Val Ala Thr Ser Gln Asp Asp Cys Tyr Lys Phe Ala Ile			
2877	tca cag tca tcc acg ggc act gtt atg gga gct gtt atc atg gag ggc	390	395	400
	Ser Gln Ser Ser Thr Gly Thr Val Met Gly Ala Val Ile Met Glu Gly			
2925	ttc tac gtt gtc ttt gat cgg gcc cga aaa cga att ggc ttt gct gtc	405	410	415
	Phe Tyr Val Val Phe Asp Arg Ala Arg Lys Arg Ile Gly Phe Ala Val			
2973	agc gct tgc cat gtg cac gat gag ttc agg acg gca gcg gtg gaa ggc	420	425	430
	Ser Ala Cys His Val His Asp Glu Phe Arg Thr Ala Ala Val Glu Gly			

Figure 13G

cct ttt gtc acc ttg gac atg gaa gac tgt ggc tac aac att cca cag	3021
Pro Phe Val Thr Leu Asp Met Glu Asp Cys Gly Tyr Asn Ile Pro Gln	
435 440 445	
<hr/>	
aca gat gag tca acc ctc atg acc ata gcc tat gtc atg gct gcc atc	3069
Thr Asp Glu Ser Thr Leu Met Thr Ile Ala Tyr Val Met Ala Ala Ile	
450 455 460 465	
<hr/>	
tgc gcc ctc ttc atg ctg cca ctc tgc ctc atg gtg tgt cag tgg cgc	3117
Cys Ala Leu Phe Met Leu Pro Leu Cys Leu Met Val Cys Gln Trp Arg	
470 475 480	
<hr/>	
tgc ctc cgc tgc ctg cgc cag cat gat gac ttt gct gat gac atc	3165
Cys Leu Arg Cys Leu Arg Gln Gln His Asp Asp Phe Ala Asp Ile	
485 490 495	
<hr/>	
tcc ctg ctg aag tga ggaggcccat gggcagaaga tagagattcc cctggaccac	3220
Ser Leu Leu Lys	
500	
<hr/>	
acctccgtgg ttcactttgg tcacaagtag gagacacaga tggcacctgt ggccagagca	3280
<hr/>	
cctcaggacc ctccccacc accaaatgcc tctgccttga tggagaagga aaaggctggc	3340
<hr/>	
aagggtgggtt ccagggactg tacctgtagg aacagaaaa gagaaagaa aagcactctg	3400
<hr/>	
ctggcgggaa tactcttggt cacctcaaat ttaagtctgg aaattctgct gcttgaaact	3460
<hr/>	

Figure 13H

tcagccctga acctttgtcc accattcctt taaattctcc aacccaaagt attcttcttt 3520
tcttagtttc agaagtactg gcatcacacg caggttacct tggcgtgtgt ccctgtggtg 3580
HindIII
ccctggcaga gaagagacca agcttgtttc cctgctggcc aaagtcagta ggagaggatg 3640
cacagtttgc tatttgcttt agagacaggg actgtataaa caagccctaac attggtgcaa 3700
agattgcctc ttgaattaaa aaaaaaact agattgacta ttatacaaa tgggggcggc 3760
tggaaaagagg agaaggagag ggagtacaaa gacaggggaat agtgggatca aagctaggaa 3820
aggcagaaac acaaccactc accagtccta gtttagacc tcattctcaa gatagcatcc 3880
catctcagaa gatgggtgtt gttttcaatg ttttcttttc tgtggttgca gcctgaccaa 3940
aagtgagatg ggaagggctt atctagccaa agagctcttt ttagctctc ttaaatgaag 4000
tgcccactaa gaagtccac ttaacacatg aatttctgcc atattaattt cattgtctct 4060
atctgaacca ccctttattc tacatatgat aggcagcact gaaatatcct aaccccctaa 4120
gctccaggtg ccctgtggga gagcaactgg actatagcag ggctgggctc tgtctcctg 4180
gtcataggct cactcttttc cccaaatctt cctctggagc ttgacagcca aggtgctaaa 4240
aggaataggt aggagacctc ttctatctaa tccttaaaag cataatgttg aacattcatt 4300

Figure 13I

caacagctga tgcctataa cccctgcctg gatttcttcc tattaggcta taagaagtag 4360
caagatcttt acataattca gagtgtttc attgccttcc taccctctct aatggccctt 4420
ccatttat tt gactaaagca tcacacagtgc gactagcat tataccaaga gtatgagaaa 4480
tacagtgc tt tatggctcta acattactgc cttcagtatc aaggctgcct ggagaaagga 4540
tggcagcctc agggcttctt tatgtcctcc accacaagag ctccctgatg aaggtcatct 4600
ttttccccta tcctgttctt cccctccccg ctccaatgg tacgtgggta cccaggctgg 4660
ttcttgggct aggtagtggg gaccaagtcc attacctcc tatcagttct agcatagtaa 4720
actacggtag cagtgttagt gggaagagct gggtttctct agtatacca ctgcataccta 4780
ctcctacctg gtcaacccgc tgcttcagg tatgggacct gctaagtgtg gaattacctg 4840
ataagggaga gggaataca aggaggccct ctggtgttcc tggcctcagc cagctgcccc 4900
caagccataa accaataaaa caagaatact gagtcagttt ttctcttcat 4960
tccactgca ctgtgtgtg ctttggctga ctgggaacac ccataacta cagagtctga 5020
caggaagact ggagactgtc cacttctagc tcggaactta ctgtgtaaat aaactttcag 5080
aactgtacc atgaagtga aatgccacat ttgtctttat aatttctacc catgttggga 5140

Figure 13J

aaaactggct ttttccagc cttttccagg gcataaaact caacccttc gatagcaagt 5200
cccatcagcc tattattttt ttaaagaaaa cttgcacttg tttttctttt tacagttact 5260
tccttcctgc cccaaaaatta taaactctaa gtgtaaaaaa aagtcttaac aacagcttct 5320
tgcttgtaaa aatatgtatt atacatctgt atttttaaat tctgctcctg aaaaatgact 5380
gtcccatctc cactcactg catttggggc ctttcccat ggtctgcatg tcttttatca 5440
ttgcaggcca gtggacagag ggagaaggga gaacaggggt cgccaacact tgtgttgctt 5500
tctgactgat cctgaacaag aaagagtaac actgaggcg cgtctcccat gcacaaactc 5560
ccaaaacact tctcctctg caagagtggg ctttccgggt ctttactggg aagcagttaa 5620
gccccctcct cacccttcc ttttttcttt ctttactcct ttggcttcaa aggattttgg 5680
aaaagaaca atatgcttta cactcathtt caatttctaa attgcaggg gatactgaaa 5740
aatacggcag gtggcctaag gctgctgtaa agttgagggg agaggaaatc ttaagattac 5800
aagataaaaa acgaatcccc taaacaaaaa gaacaataga actggtcttc cattttgcca 5860
cctttcctgt tcatgacagc tactaacctg gagacagtaa catttcatta accaaagaaa 5920
gtgggtcacc tgacctctga agagctgagt actcaggcca ctccaatcac cctacaagat 5980

Figure 13K

gccaaaggagg tcccaggaag tccagctcct taaactgacg ctagtcaata aacctggggca 6040
agtgaggcaa gagaaatgag gaagaatcca tctgtgaggt gacaggcacg gatgaaagac 6100
aaagacgga aagagtatca aaggcagaaa ggagatcatt tagttgggtc tgaaggaaa 6160
agtntttgct atccgacatg tactgctagt wcctgtaagc attttaggtc ccagaatgga 6220
aaaaaaaaatc aagctatngg ttataataata atgnnnnnnnn nnnnnnnnnn nntcgagcat 6280
gcatctagag ggccttattc tatagtgtca cctaaatgct agagctcgct gatcagcctc 6340
gactgtgcct tctagttgcc agccatctgt tgtttgcccc tccccgtgc cttccttgac 6400
cctggaagggt gccactcca ctgtcccttc ctaataaaat gaggaattg catcgcatcg 6460
tctgagtagg tgtcattcta ttctggggggg tgggggtgggg caggacagca agggggaggga 6520
ttgggaagac aatagcaggc atgctgggga tgcggtgggc tctatggctt ctgaggcgga 6580
aagaaccagc tggggctcta gggggtatcc ccacgcgccc tgtagcggcg cattaagcgc 6640
ggcgggtgtg gtggttacgc gcagcgtgac cgctacactt gccagcgccc tagcgcccc 6700
tcctttcgct ttcttccctt cctttctcgc cacgttcgcc ggctttcccc gtcaagctct 6760
aaatcggggc atccctttag ggttccgatt tagtgcttta cggcacctcg accccaaaa 6820

Figure 13I

acttgattag ggtgatggtt cacgtagtgg gccatcgccc tgatagacgg tttttcgccc 6880
tttgacgttg gagtccacgt tctttaatag tggactcttg ttccaaactg gaacaacact 6940
caaccctatc tcggtctatt cttttgattt ataaggatt ttggggattt cggcctattg 7000
gttaaaaaat gagctgattt acaaaaaatt taacgcgaat tctagagccc cgccgccgga 7060
cgaactaaac ctgactacgg catctctgcc cttctctgc ggggcagtgc atgtaatccc 7120
ttcagttggt tggtaacaact tgccaactgg gccctgttcc acatgtgaca cgggggggga 7180
ccaaacacaa aggggttctc tgactgtagt tgacatcctt ataaatggat gtgcacattt 7240
gccaacactg agtggcttc atcctggagc agactttgca gtctgtggac tgcaaacacaa 7300
cattgccttt atgtgtaact cttggctgaa gctcttacac caatgctggg ggacatgtac 7360
ctcccagggg cccaggaaga ctacgggagg ctacaccaac gtcaatcaga ggggcctgtg 7420
tagctaccga taagcggacc ctcaagaggg cattagcaat agtgtttata agggccctt 7480
HpaI
gttaacccta aacgggtagc atatgcttcc cgggtagtag tatactat ccagactaac 7540
cctaattcaa tagcatatgt tacccaacgg gaagcatatg ctatcgaatt agggttagta 7600
EcoRV
aaagggctct aaggaacagc gatattctcc acccatgag ctgtcacggt ttatttaca 7660

Figure 13M

tggggtcagg attccacgag ggtagtgaa catttttagtc acaagggcag tggctgaaga 7720
tcaaggagcg ggcagtgaac tctcctgaat cttcgcctgc ttcttcattc tccttcgttt 7780
agctaataga ataactgctg agttgtgaac agtaagggtg atgtgagggtg ctgaaaaaca 7840
aggtttcagg tgacgcccc agaataaaat ttggacgggg ggttcagtgg tggcattgtg 7900
ctatgacacc aatataaacc tcacaaaacc cttgggcaat aaatactagt gtaggaatga 7960
aacattctga atatctttaa caatagaaat ccatgggggtg gggacaagcc gtaaagactg 8020
gatgtccatc tcacacgaat ttatggctat gggcaaacaca taatccctagt gcaatatgat 8080
actgggggta ttaagatgtg tcccaggcag ggaccaagac aggtgaacca tgttgttaca 8140
ctctatttgt aacaagggga aagagagtgg acgccgacag cagcggactc cactggttgt 8200
ctctaacacc cccgaaaaat aaacggggct ccacgccaat ggggcccata aacaaagaca 8260
agtggccact cttttttttg aaatttgtga gtggggggcac gcgtcagccc ccacacgccg 8320
ccctgcggtt ttggactgta aaataagggt gtaataactt ggctgattgt aaccccgcta 8380
accactgcgg tcaaaaccact tgcccacaaa accactaatg gcaccccggtg gaatacctgc 8440
ataagtaggt gggcgggcca agatagggc gcgattgctg cgatctggag gacaaattac 8500

Figure 13N

```

acacacttgc gcctgagcgc caagcacagg gttgttggtc ctcatattca cgaggtcgct 8560
gagagcacgg tgggctaattg ttgccatggg tagcatatac taccctaaata tctggatagc 8620
atatgctatc ctaatctata tctgggtagc ataggctatc ctaatctata tctgggtagc 8680
atatgctatc ctaatctata tctgggtagt atatgctatc ctaatttata tctgggtagc 8740
ataggctatc ctaatctata tctgggtagc atatgctatc ctaatctata tctgggtagt 8800
atatgctatc ctaatctgta tccgggtagc atatgctatc ctaatagaga ttagggtagt 8860
atatgctatc ctaatttata tctgggtagc atatactacc caaatatctg gatagcatat 8920
gctatccctaa tctatatctg ggtagcatat gctatccctaa tctatatctg ggtagcatag 8980
gctatccctaa tctatatctg ggtagcatat gctatccctaa tctatatctg ggtagtatat 9040
gctatccctaa tttatatctg ggtagcatag gctatccctaa tctatatctg ggtagcatat 9100
gctatccctaa tctatatctg ggtagtatat gctatccctaa tctgtatccg ggtagcatat 9160
gctatccctca tgcatatata gtcagcatat gataccagc agtagagtgg gtagtctatc 9220
ctttgcatat gccgccacct cccaaggggg cgtgaatttt cgctgcttgt ccttttccctg 9280
catgctggtt gctcccatc ttagtgaat taaaggaggc caggctaaag ccgtcgcatg 9340

```

oriP

↓

Figure 130

tctgattgct caccaggtaa atgtcgctaa tgttttccaa cgcgagaagg tgttgagcgc 9400
ggagctgagt gacgtgacaa catgggtatg cccaattgcc ccatgttggg aggacgaaaa 9460
tggtgacaag acagatggcc agaaatacac caacagcacg catgatgtct actgggggatt 9520
tattcttttag tgcgggggaa tacacggctt ttaatacgat tgagggcgctc tcctaacaag 9580
ttacatcact cctgcccttc ctacccctca tctccatcac ctcttcac tcctcatct 9640
ccgtcatcac cctccgcggc agcccccttc accatagggtg gaaaccaggg aggcaaatct 9700
actccatcgt caaagctgca cacagtcacc ctgatatgc aggtaggagc gggcttctgc 9760
ataacaaggc ccttaatcgc atccttcaaa acctcagcaa atatatgagt ttgtaaaaag 9820
accatgaaat aacagacaaat ggactccctt agcggggccag gttgtgggcc gggctccagg 9880
gccattccaa aggggagacg actcaatggt gtaagacgac attgtggaat agcaagggca 9940
gttcctcgc ttaggttgta aaggagggtc ttactacctc catatacgaa cacaccggcg 10000
acccaagtcc cttcgtcggc agtcctttct acgtgactcc tagccaggag agctcttaa 10060
ccttctgcaa tgttctcaaa ttctcgggttg gaacctcctt gaccacgatg ctttccaaac 10120
cacctcctt ttttgcgcct gcctccatca ccctgacccc ggggtccagt gcttgggcct 10180

Figure 13P

tctcctgggt catctgagg gccctgctct atcgctccg ggggcacgtc aggtcacca 10240
tctggggcac cttcttggtg gattcaaaa taatcggctt ccctacagg gtggaaaaat 10300
ggcctctac ctggagggg cctgcgcgt ggagaccgg atgatgatga ctgactactg 10360
ggactcctgg gcctcttttc tccacgtcca cgacctctcc ccctggctct ttcacgactt 10420
ccccccctgg ctctttcacg tctctatacc cggcggcctc cactacctcc tcgaccccg 10480
cctccactac ctccctgacc ccggcctcca ctgcctcctc gaccccgcc tccacctcct 10540
gctcctgccc ctctgctcc tgccccctct cctgctcctg cccctcctgc ccctcctgct 10600
cctgccccctc ctgccccctcc tgctcctgcc cctcctgccc ctctgctcc tgccccctct 10660
gccccctcctc ctgctcctgc cctcctgccc cctcctcctg ctctgcccc tcttgcccc 10720
cctgctcctg cccctcctgc cctcctgct cctgccccctc ctgccccctc tgctcctgcc 10780
cctcctgctc ctgccccctcc tgctcctgcc cctcctgctc ctgccccctc tgccccctct 10840
gccccctcctc ctgctcctgc cctcctgct cctgccccctc ctgccccctc tgccccctct 10900
gctcctgccc ctctcctgc tctgccccct cctgccccctc ctgccccctc tctgctcct 10960
gccccctcctg cccctcctcc tgctcctgcc cctcctcctg ctctgcccc tctgcccc 11020

Figure 13Q

cctgcccctc ctctgtctcc tgccccctct tgcctctctc ctgtctctgc cctctctctc 11080
gctcctgccc ctcttgcccc tcttgcccc ctctctgctc ctgccccctc tctgtctctc 11140
gccccctctg cccctctctgc ccctctctgcc cctctctctg ctcttgcccc tctctctgct 11200
cctgccccct ctgtctctctgc ccctccccgt cctgtctctg ctctgttctc accgtgggtc 11260
cctttgcagc caatgcaact tggacgtttt tgggggtctc ggacaccatc tctatgtctc 11320
ggccctgctc ctgagccgcc cggggctctc ggtcttccgc ctctctgtcc tcgtcctctc 11380
ccccgtcctc gtccatgggt atcaccccc ctctcttgag gtccactgcc gccggagcct 11440
tctgggtccag atgtgtctcc ctctctctc aggccatttc caggctctgt acctggcccc 11500
tcgtcagaca tgattcacac taaaagagat caatagacat ctttattaga cgacgtcag 11560
tgaatacagg gagtgcagac tcttgcccc tccaaacagc cccccacct catccccctc 11620
atgggtcgctg tcagacagat ccagggtctga aaattcccca tcctccgaac catcctcgtc 11680
ctcatcacca attactcgca gcccggaata ctcccgctga acatcctcaa gatttgcgtc 11740
ctgagcctca agccaggcct caaattcctc gtcccccttt ttgctggacg gtagggatgg 11800
ggattctcgg gaccctctct ctctctctc aaggtcacca gacagagatg ctactggggc 11860

Figure 13R

ClalI

aacggaagaa aagctgggtg cggcctgtga ggatcagctt atcgatgata agctgtcaaa 11920
catgagaatt cttgaagacg aaagggcctc gtgatacgcc tatttttata ggttaatgtc 11980
atgataataa tggtttctta gacgtcaggt ggcacttttc ggggaaatgt gcgcggaacc 12040
cctatttggt tatttttcta aatacatcca aatatgtatc cgctcatgag acaataaacc 12100
tgataaaatgc ttcaataata ttgaaaaagg aagagtatga gtattcaaca ttcccggtgc 12160
gcccttattc ccttttttgc ggcattttgc ctccctgttt ttgctcacc agaaacgctg 12220
gtgaaaagtaa aagatgctga agatcagttg ggtgcacgag tgggttacat cgaactggat 12280
ctcaacacgc gtaagatcct tgagagtttt cgccccgaag aacgttttcc aatgatgagc 12340
actttttaag ttctgctatg tggcgcggtta ttatcccgtg ttgacgccgg gcaagagcaa 12400
ctcggtcgcc gcatacacta ttctcagaat gacttggttg agtactcacc agtcacagaa 12460
aagcatctta cggatggcat gacagtaaga gaattatgca gtgctgccat aacctgagt 12520
gataacactg cggccaactt acttctgaca acgatcggag gaccgaagga gctaaccgct 12580
tttttgcaca acatggggga tcatgtaact cgccttgatc gttgggaacc ggagctgaat 12640
gaagccatac caaacgacga gcgtgacacc acgatgcctg cagcaatggc aacaacgttg 12700

Figure 13S

cgcaaaactat taactggcga actacttact ctagcttccc ggcaacaatt aatagactgg 12760
atggaggcgg ataaagtgc aggaccactt ctgcgctcgg cccttccggc tggctggttt 12820
attgctgata aatctggagc cggtgagcgt gggctctcgg gtatcattgc agcactgggg 12880
ccagatggta agccctcccg tatcgtagtt atctacacga cggggagtcg ggcaactatg 12940
gatgaacgaa atagacagat cgctgagata ggtgcctcac tgattaagca ttggtaactg 13000
tcagaccaag tttactcata tatacttttag attgatttaa aacttcatth ttaatttaaa 13060
aggatctagg tgaagatcct ttttgataat ctcatgacca aaatccctta acgtgagttt 13120
tcgttccact gagcgtcaga ccccgtagaa aagatcaaaag gatcttcttg agatcctttt 13180
tttctgcgcg taatctgctg cttgcaaaca aaaaaaacac cgctaccagc ggtgggttgt 13240
ttgccggatc aagagctacc aactcttttt ccgaaggtaa ctggcttcag cagagcgcag 13300
ataccaaata ctgtccttct agtgtagccg tagttaggcc accacttcaa gaactctgta 13360
gcaccgccta catacctcg cctgctaata ctgttaccag tggctgctgc cagtggcgat 13420
aagtcgtgtc ttaccgggtt ggactcaaga cgatagttac cggataaggc gcagcgggtcg 13480
ggctgaacgg ggggttcgtg cacacagccc agcttgagc gaacgacctg caccgaactg 13540

Figure 13T

agatacctac agcgtgagct atgagaaagc gccacgcttc ccgaaggag aaaggcggac 13600
 aggtatccgg taagcggcag ggtcgggaaca ggagagcgca cgagggagct tccaggggga 13660
 aacgcctggt atctttatag tcctgtcggg ttctgccacc tctgacttga gcgtcgattt 13720
 ttgtgatgct cgtcaggggg gcggagccta tggaaaaacg ccagcaacgc ggccttttta 13780
 cggttccctgg ccttttgctg cgccgcgtgc ggctgctgga gatggcggac gcgatggata 13840
 tgttctgcca agggttgggt tgcgcattca cagttctccg caagaattga ttggctccaa 13900
 ttcttggagt ggtgaatccg ttagcgaggt gccgcgggt tccattcagg tcgaggttggc 13960
 ccggctccat gcaccgcgac gcaacgcggg gaggcagaca aggtataggg cggcgcctac 14020
 aatcccatgcc aaccggttcc atgtgctcgc cgaggcggca taaatcgccg tgacgatcag 14080
 cggtcragt atcgaagtta ggctggtaag agccgcgagc gatccttgaa gctgtccctg 14140
 atggtcgtca tctacctgcc tggacagcat ggcctgcaac gcgggcatcc cgatgccgcc 14200
 ggaagcgaga agaatacataa tggggaaggc catccagcct cgcgtcgcga acgccagcaa 14260
 gacgtagccc agcgcgtcgg ccgccatgcc ctgcttcac cccgtggccc gttgctcgcg 14320
tttgctggcg gtgtcccccg aagaaatata tttgcatgtc ttagttcta tgatgacaca 14380

NruI

Figure 13U

aaccccgccc agcgtcttgt cattggcgaa ttcgaacacg cagatgcagt cggggcggcg 14440
cgggtcccagg tccacttcgc atattaaggt gacgcgtgtg gcctcgaaca ccgagcgacc 14500
ctgcagcgac ccgcttaaca gcgtcaacag cgtgccgcag atcccgggca atgagatatg 14560
aaaaagcctg aactcaccgc gacgtctgtc gagaagtttc tgatcgaaaa gttcgacagc 14620
gtctccgacc tgatgcagct ctcgaggggc gaagaatctc gtgctttcag cttcgatgta 14680
ggagggcgtg gatatgtcct gcgggtaaat agctgcgccg atggtttcta caaagatcgt 14740
tagtggggatc ggcactttgc atcggccgcg ctcccgatt ccggaagtgc ttgacattgg 14800
ggaattcagc gagagccctga cctattgcat ctcccgccgt gcacaggggtg tcacgttgca 14860
agacctgcct gaaaccgaac tgcccgcctgt tctgcagccg gtcgcggagg ccattggatgc 14920
gacgcgtgcg gccgatctta gccagacgag cgggttcggc ccatcggac cgcaagggaat 14980
cgggtcaatac actacatggc gtgatttcat atgcgcgatt gctgatcccc atgtgtatca 15040
ctggcacaact gtgatggacg acaccgtcag tgcgtccgtc gcgcaggctc tcgatgagct 15100
gatgctttgg gccgaggact gccccgaagt ccggcacctc gtgcacgcgg atttcggctc 15160
caacaatgtc ctgacggaca atggccgcat aacagcggtc attgactgga gcgaggcgat 15220

Figure 13V

gttcgggggat tccaatacag aggtcgcgcaa catcttcttc tggaggccgt ggttggcggg 15280
tatggagcag cagacgcgct acttcgagcg gaggcataccg gagcttgca gatacgcgcg 15340
gctccgggcg tatatgctcc gcatagggtct tgaccaactc tatcagagct tggttgacgg 15400
caatttcgat gatgcagctt gggcgacggg tcgatgcgac gcaatcgtcc gatccggagc 15460
cgggactgtc gggcgctacac aaatcgcccg cagaagcgcg gccgtctgga ccgatggctg 15520
tgtagaagta ctcgccgata gtggaaacgg gagatggggg aggctaactg aaacacggaa 15580
ggagacaata ccggaaggaa cccgcgctat gacggcaata aaaagacaga ataaaacgca 15640
cgggtgttgg gtcgttttgtt cataaacgcg gggttcggtc ccagggctgg cactctgtcg 15700
ataccccacc gagaccccat tggggccaat acgccgcgt ttcttcttt tcccccccc 15760
accccccaag ttcgggtgaa ggcccagggc tcgagccaa cgtcggggcg gcaggccctg 15820
ccatagccac tggccccctg ggttagggac ggggtcccc atgggggaatg gtttatggtt 15880
cgtgggggtt attattttgg gcgttgctg gggcttggtc cagactgga ctgagcagac 15940
agaccccatg tttttggatg gcctgggcat ggaccgcatg tactggcgcg acacgaacac 16000
cgggcgtctg tggctgcaa acacccccga cccccaaaa ccaccgcgcg gatttctggc 16060

Figure 13W

SalI

gtgccaaagct agtcgaccaa
▲

16080

CTGTTGGGCTCGCGGTTGAGGACAAACTCTTCGCGGTCTTTCCAGTACTCTTGGATCGGAAAC
 CCGTCGGCCTCCGAACGGTACTCCGCCACCGAGGGACCTGAGCGAGTCCGCATCGACCGGAT
 CGGAAAACCTCTCGACTGTTGGGGTGAGTACTCCCTCTCAAAAGCGGGCATGACTTCTGCGCT
 AAGATTGTCAGTTTCCAAAAACGAGGAGGATTTGATATTACCTGGCCCGCGGTGATGCCTTT
 GAGGGTGGCCGCGTCCATCTGGTCAGAAAAGACAATCTTTTTGTTGTCAAGCTTGAGGTGTGG
 CAGGCTTGAGATCTGGCCATACACTTGAGTGACAATGACATCCACTTTGCCTTTCTCTCCACAG
 GTGTCCACTCCCAGGTCCAACCTGCAGGTCGACTCTAGACCC

FIG. 14A

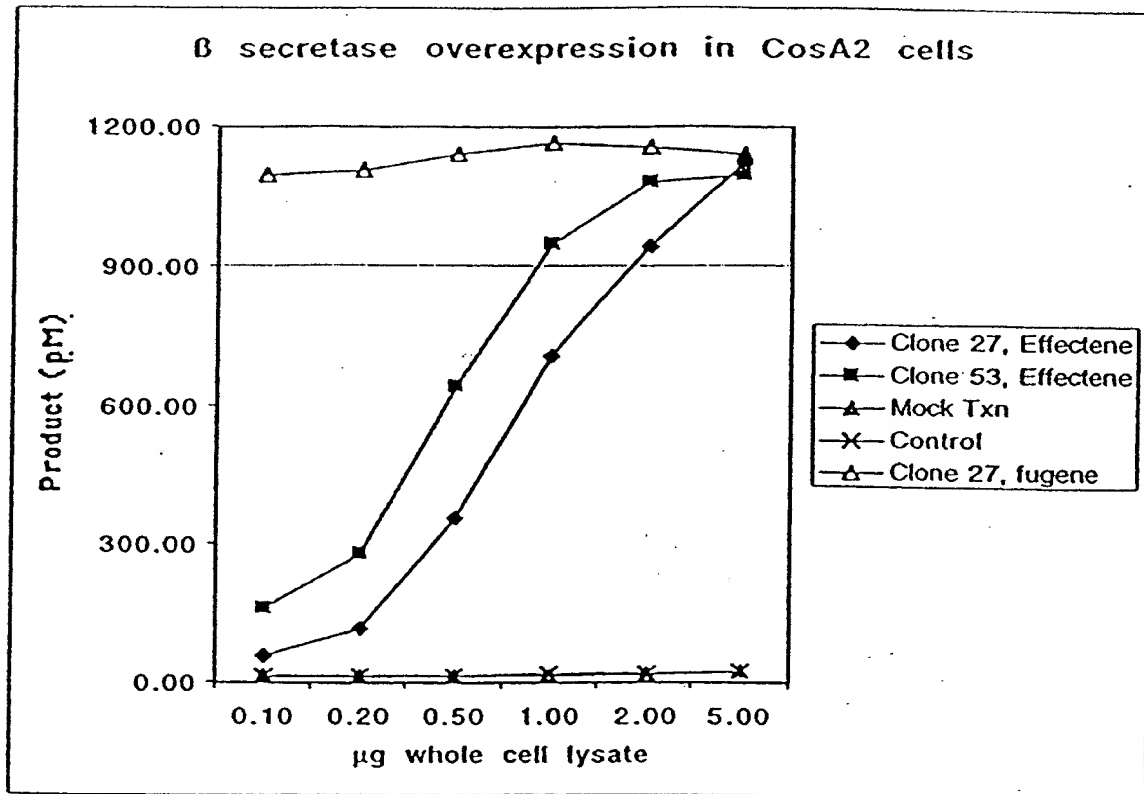


FIG. 14B

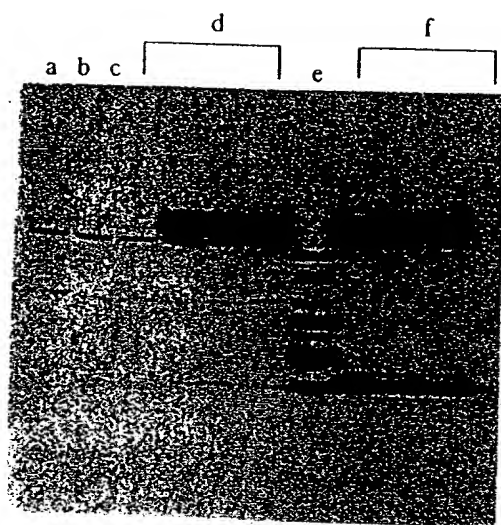


FIG. 15A

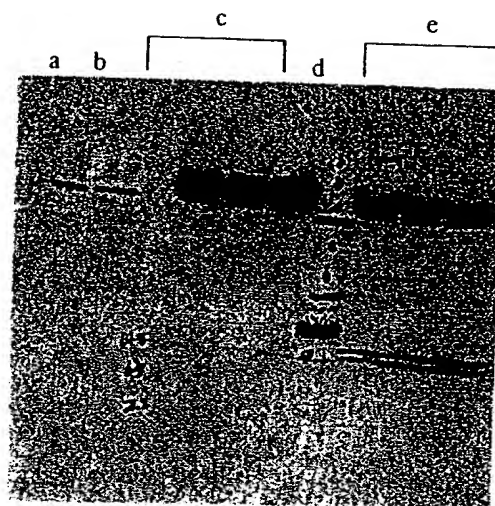


FIG. 15B

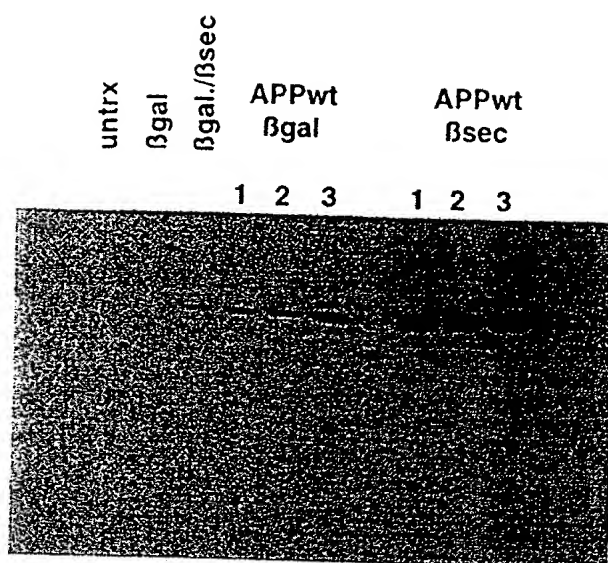


FIG. 16A

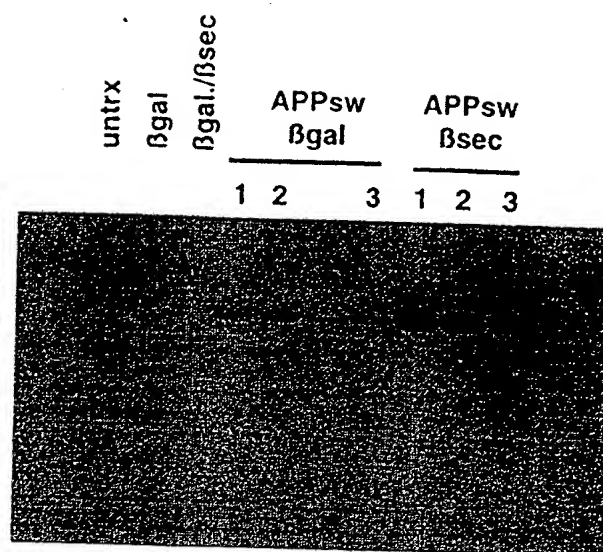


FIG. 16B

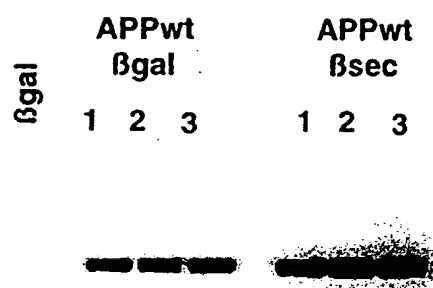
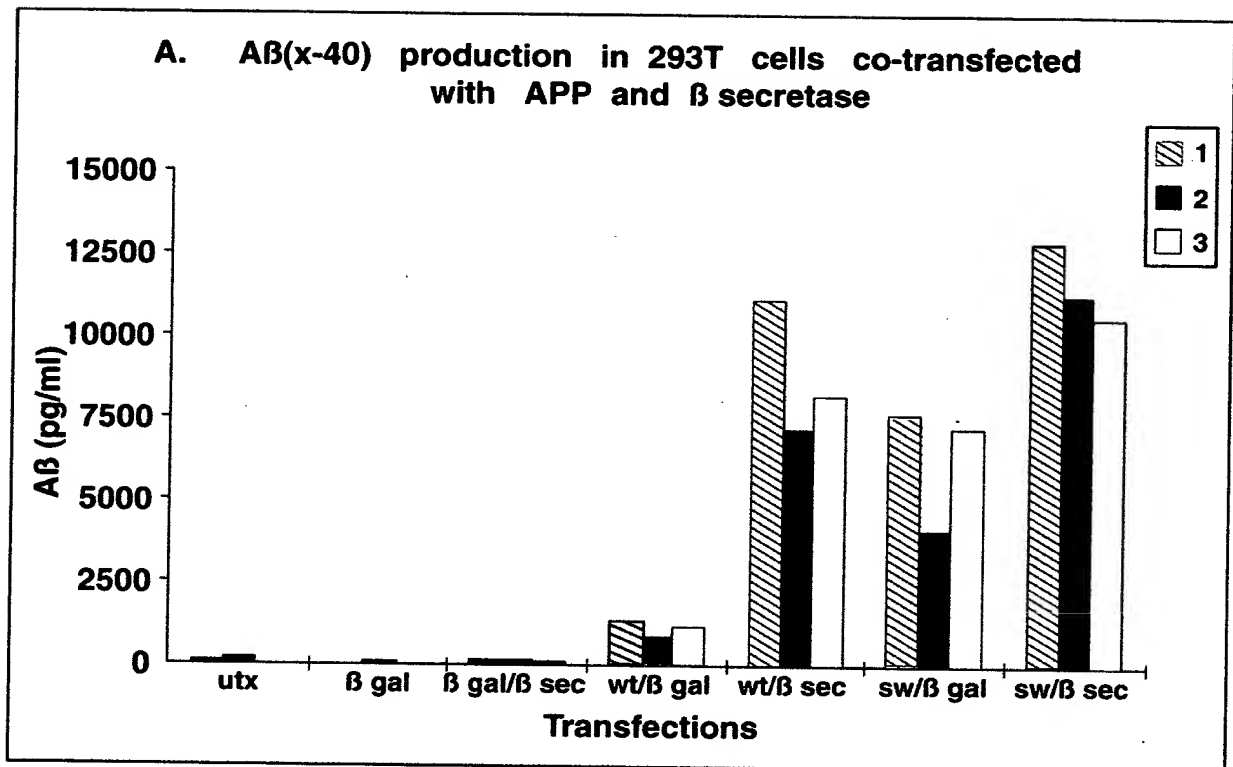


FIG. 17A



FIG. 17B

**Fig. 18**

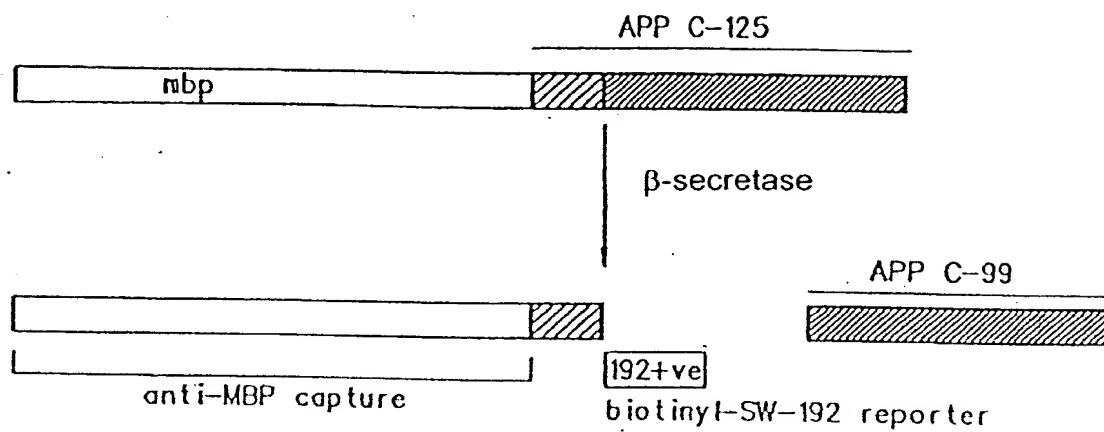


FIG. 19A

Wild-Type SequenceVal-Lys-Met-Asp...
Swedish SequenceVal-Asn-Leu-Asp...

FIG. 19B

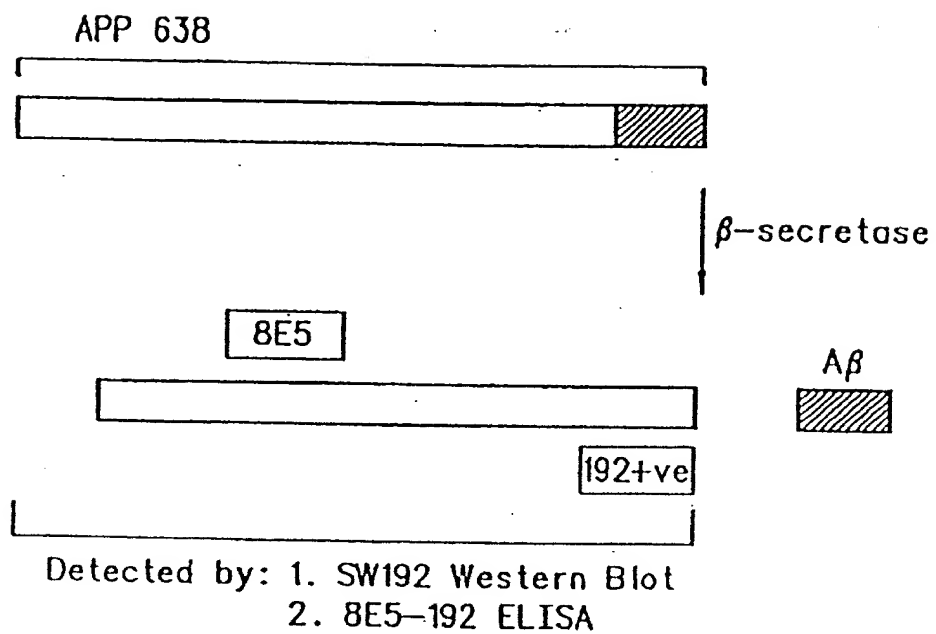


FIG. 20

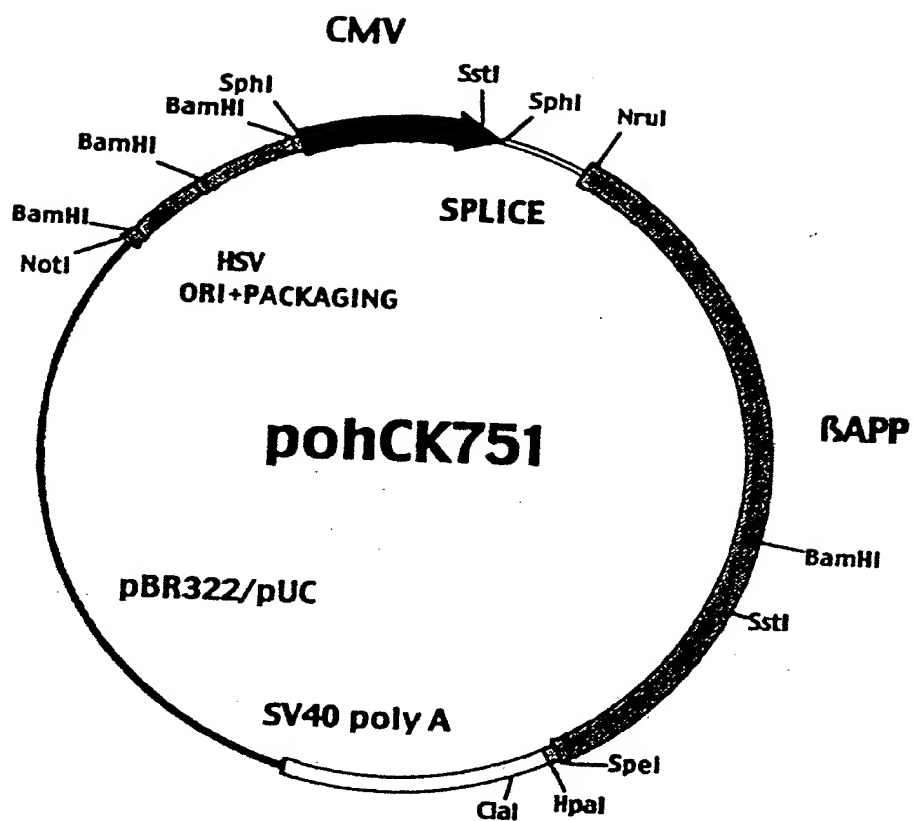


FIG. 21